(51) **International Patent Classification[7]:** G10L 21/02

(21) **International Application Number:** PCT/US02/01113

(22) **International Filing Date:** 17 January 2002 (17.01.2002)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**
60/264,924   30 January 2001 (30.01.2001)   US

(71) **Applicant** *(for all designated States except US)*: THOMSON LICENSING S.A. [FR/FR]; 46, quai A. Le Galo, 92648 Boulogne Cedex (FR).

(72) **Inventors; and**
(75) **Inventors/Applicants** *(for US only)*: PARRA, Lucas,

Cristobal [DE/US]; 160 Cabrini Blvd. #131, New York, NY 10033 (US). ALVINO, Christopher [US/US]; 2128 Saw Mill Lane, Allenwood, NJ 08720-0189 (US). SPENCE, Clay, Douglas [US/US]; 136 Cranbury Road, Princeton Junction, NJ 08550 (US). FANCOURT, Craig, Langdale [US/US]; 5531 Town Court South, Lawrenceville, NJ 08648 (US).

(74) **Agents:** TRIPOLI, Joseph, S. et al.; Thomson multimedia Licensing Inc., P.O. Box 5312, Princeton, NJ 08540 (US).

(81) **Designated States** *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(54) **Title:** GEOMETRIC SOURCE SEPARATION SIGNAL PROCESSING TECHNIQUE
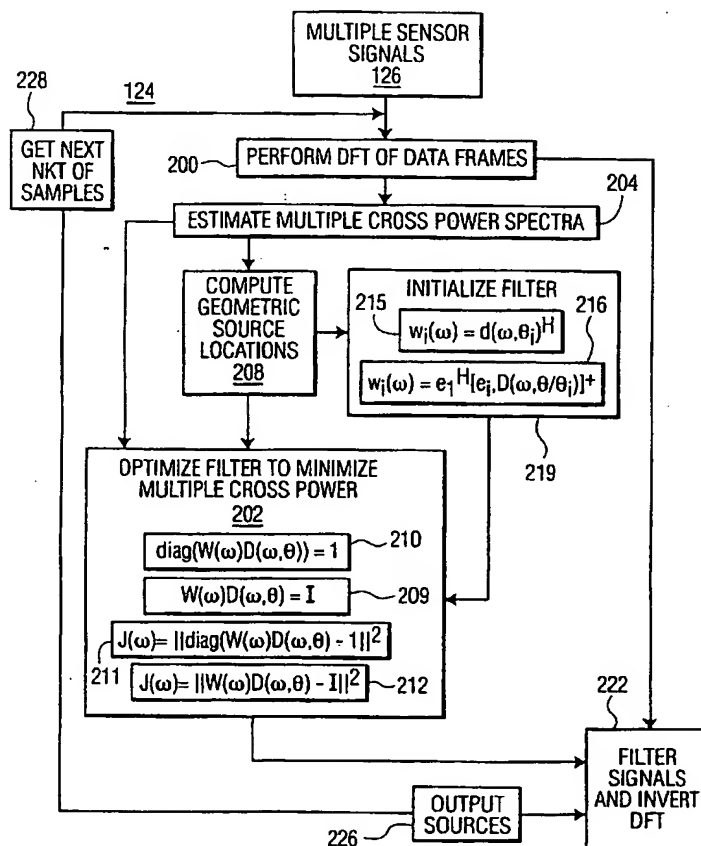
(57) **Abstract:** A computer system (108) that processes mixtures of signals, such as speech and noise sources derived from multiple simultaneous microphone recordings, in order to separate them into their underlying sources. A source separation routine (124) optimizes a filter structure by minimizing cross powers of multiple output channels while enforcing geometric constraints on the filter response. The geometric constraints (209, 210, 215, 217) enforce desired responses for given locations of the underlying sources, based on the assumption that the sources are localized in space.

WO 02/061732 A1

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv)) for US only*

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

GEOMETRIC SOURCE SEPARATION SIGNAL PROCESSING TECHNIQUE

1. FIELD OF THE INVENTION

This invention pertains generally to signal processing and more specifically to a system for performing mixed signal separation using geometric information and adaptive beamforming techniques.

2. BACKGROUND OF THE INVENTION

Blind source separation refers to the process of separating a composite signal into its original component signals without prior knowledge of their characteristics. This process is useful in speech recognition, multipath channel identification and equalization, improvement of the signal to interference ratio (SIR) of acoustic recordings, in surveillance applications and in the operation of hearing aids.

Blind source separation of broad band signals in a multipath environment remains a difficult problem which has a number of ambiguities. Increasing the number of sensors allows improved performance but leads to ambiguities in the choice of the separating filters. There are in theory multiple filters that invert the responses in a room because there are multiple projections from the space containing microphone signals into the smaller space of signal sources. These multiple filters represent remaining degrees of freedom in terms of a sensor array response.

The consistent assignment of signal contributions to different source channels across different frequencies creates a frequency permutation problem. This problem is inherent to all source separation algorithms including time domain algorithms unless the algorithm simultaneously considers different frequency bands. An estimation of such polyspectral properties is particularly difficult for nonstationary signals such as speech and the resulting algorithms are computationally expensive.

The basic source separation problem is described by assuming the existence of M

uncorrelated, time varying source signals source signals:

$$s(t) \in \mathbf{R}^M$$

5

where the sources s(t) originate from different spatial locations. A number of sensors N

(where $N \geq M$) detect time varying signals:

$$x(t) \in \mathbf{R}^N.$$

10

In a multipath environment each source j couples with sensor i through a linear transfer

function $A_{ij}(\tau)$, representing the impulse response of the corresponding source to sensor path

such that:

15 $$x_i(t) = \sum_{j=1}^{M} \sum_{\tau=0}^{P-1} A_{ij}(\tau) s_j(t-\tau).$$

This equation can be rewritten using matrix notation (denoting the convolutions by *):

$$x(t) = A(t) * s(t).$$

20

After applying the discrete time Fourier transform, this equation may be rewritten as:

25 $$x(\omega) = A(\omega)s(\omega).$$

The goal of convolutive source separation is to find finite impulse response (FIR) filters

$W_{ij}(\tau)$ that invert the effect of the convolutive mixing $A(\tau)$. This is equivalent to producing

$$y(\omega) = W(\omega)x(\omega)$$

30

that correspond to the original sources s(t).

Different criteria for convolutive separation have been proposed, for example as discussed by H.-L. N. Thi and C. Jutten in "BLIND SOURCE SEPARATION FOR CONVOLUTIVE MIXTURES", published in Signal Processing, vol. 45, no. 2, pp. 209-229 (1995). A two channel example is disclosed in U.S. Patent No. 5,208,786 entitled MULTI-

5   CHANNEL SIGNAL SEPARATION, issued to Weinstein et al. The '786 patent models each channel as a multi-input-multi-output (MIMO) linear time invariant system. The input source signals are separated and recovered by requiring that the reconstructed signals be statistically uncorrelated. However, the decorrelation condition is insufficient to uniquely solve the problem unless one assumes that the unknown channel is a 2 X 2 MIMO finite impulse

10   response filter.

All convolutive separation criteria can be derived from the assumption of statistical independence of the unknown signals, usually limited to pairwise independence of the source signals. Pairwise independence implies that all cross-moments can be factored, thereby

15   creating a number of necessary conditions for the model signal sources,

$\forall\, t, n, m, \tau, i \neq j$:

$$E\left[y_i^n(t)y_j^m(t+\tau)\right] = E\left[y_i^n(t)\right] E\left[y_j^m(t+\tau)\right]. \qquad \text{(Equation I)}$$

20   Convolutive separation requires that these conditions by satisfied for multiple delays τ which correspond to the delays of the filter taps of $W(\tau)$. For stationary signals higher order criteria (multiple n, m) are required. For non-stationary signals such as speech multiple t can be used and multiple decorrelation (n = m = 1) is sufficient.

25   When using an independence criterion there remains both a permutation and scaling ambiguity. In the convolutive case the scaling ambiguity applies to each frequency group or bin resulting in a convolutive ambiguity for each source signal in the time domain. Any delayed or convolved versions of independent signals remain independent. For the independent frequency domain

30

$$E\left[y_i^n(\omega)y_j^m(\omega)\right] = E\left[y_i^n(\omega)\right] E\left[y_j^m(\omega)\right],$$

there is a permutation ambiguity per each frequency for all orders n and m. For each frequency the independent frequency domain is therefore also satisfied by arbitrary scaling and assignment of indices i, j to the model sources

5      $$W(\omega) \, A(\omega) = P(\omega)S(\omega),\qquad \text{(Equation II)}$$

where $P(\omega)$ represents an arbitrary permutation matrix and $S(\omega)$ an arbitrary diagonal scaling matrix for each frequency. This creates the problem that contributions of a given signal source may not be assigned consistently to a single model source for different frequency bins.

10   A given model source will therefore have contributions from different actual sources. The problem is more severe with an increasing number of channels as the possible number of permutations increases.

This problem has often been considered an artifact of the frequency domain

15   formulation of the separation criteria since the separation task is decoupled into independent separation tasks per frequency bin. For $n = m = 1$ this ambiguity also applies to the time domain independence criteria set forth in Equation I. Even for higher orders the time domain criteria does not guarantee correct permutations.

20   Some source separation work in the past simply ignored the problem. Others have proposed a number of solutions such as continuity in the spectra of the model sources, or the fact that the different frequency bins are often co-modulated. A rigorous way of capturing these statistical properties of multiple frequency contributions are polyspectra. However, in practice it is difficult to obtain robust statistics at multiple frequencies, in

25   particular for non-stationary signals such as speech. In addition, the algorithms that consider combinations of frequencies are inherently computationally very demanding. Smoothness constraints on the filter coefficients in the frequency domain have also been proposed, as for example in U. S. Patent No. 6,167,417 entitled CONVOLUTIVE BLIND SOURCE SEPARATION USING A MULTIPLE DECORRELATION METHOD, issued to Parra et al.

30   This is equivalent to constraining the length of the filter as compared to the size of the analysis window. However, this limitation on the filter size may not always be reasonable as rather long filters are required in strongly reverberant environments.

In theory only N sensors are needed to separate $M = N$ sources. In practice, however, one may want to use more microphones ($N > M$) to improve the performance of a real system. Ignoring the permutation and scaling ambiguities, Equation II reads

$W(\omega)A(\omega) = I$, where I represents the identity matrix. For a given $A(\omega)$ there is a $N - M$

5   dimensional linear space of solutions $W(\omega)$, indicating that there are additional degrees of freedom when shaping a beam pattern represented by the filters $W(\omega)$.

In conventional geometric and adaptive beamforming information such as microphone position and source location is often utilized. Geometric assumptions can be incorporated

10   and implemented as linear constraints to the filter coefficients. In a multiple sidelobe canceler, for example, the response of one of the channels (channel i) is kept constant, which can be expressed as $\mathbf{w}(\omega)\mathbf{e}_i = constant$. The elements of the row vector $\mathbf{w}(\omega) \in \mathbf{C}^N$ are the filter elements to be applied to each microphone, and $\mathbf{e}_i$ is the ith column of the identity matrix. This is similar to the normalization condition imposed on the diagonal terms of W

15   that is conventionally applied in blind separation algorithms. Rather than constraining a channel one can also constrain the response of a beamformer for a particular orientation.

If the locations and response characteristics of each microphone is known, one can compute the free field response of a set of microphones and associated beam forming filters

20    $\mathbf{w}(\omega)$. For a position $\mathbf{q}$, the phase and magnitude response is given by

$$r(\omega,\mathbf{q}) = \mathbf{w}(\omega)\mathbf{d}(\omega,\mathbf{q}),$$

where $\mathbf{d}(\omega,\mathbf{q}) \in \mathbf{C}^N$ represents the phase and magnitude response of the N microphones for a

25   source located at $\mathbf{q}$. For a linear array with omnidirectional microphones and a far field source (much beyond the array aperture squared over the wavelength of interest) the microphone response depends approximately only on the angle $\theta = \theta(\mathbf{q})$ between the source and the linear array,

30      $$\mathbf{d}(\omega,\mathbf{q}) = \mathbf{d}(\omega,\theta) = e^{j\omega(p_i/c)sin(\theta)} ,$$

6

where p_i is the position of the ith microphone on the linear array and c is the wave propagation speed.

Constraining the response to a particular orientation is simply expressed by the linear constraint on $\mathbf{w}(\omega)$ such that $r(\omega,\theta) = \mathbf{w}(\omega)\mathbf{d}(\omega,\theta) = constant$. This concept is used in the linearly constrained minimum variance (LCMV) algorithm and is also the underlying idea for generalized sidelobe cancelling. In order to obtain a robust beam it has also been suggested to require a smooth response around a desired orientation. In summary, all of these conditions or a combination of them can be expressed as linear constraints on $\mathbf{w}(\omega)$.

Most adaptive beamforming algorithms consider power as their main criteria for optimization. Sometimes power is minimized such as in noise or sidelobe cancelling in order to adaptively minimize the response at the orientation of the interfering signals. Sometimes power is maximized, as in matched filter approaches, to maximize the response of interest. As a result, these algorithms often perform suboptimally when there is cross talk from other sources.

In second order source separation methods rather than considering the power of an individual beam $\mathbf{w}(\omega) \in \mathbf{C}^{1 \times N}$ and an individual channel $y(t) \in \mathbf{R}^1$, one can consider powers and cross powers of multiple beams $\mathbf{W}(\omega) \in \mathbf{C}^{M \times N}$ and their corresponding outputs $y(t) \in \mathbf{R}^M$. In the frequency domain these multiple beams and outputs correspond to the cross power spectra $\mathbf{R}_{yy}(\tau,\omega)$. Second order blind source separation of nonstationary signals minimizes cross powers across multiple times. Off diagonal elements of the matrix $\mathbf{R}_{yy}(\tau,\omega)$ are minimized in second order separation rather than the diagonal terms as is the case in conventional adaptive beamforming. Strict one channel power criteria has a serious cross talk or leakage problem when multiple sources are simultaneously active, especially in reverberant environments.

SUMMARY OF THE INVENTION

The present invention addresses the need to separate the source of nonstationary broad band signals in a multipath environment. Ambiguities in source separation are addressed by
5   adding prior information such as microphone position and by adding the assumption that the sources are localized in space. By replacing the power criteria of many adaptive beamforming algorithms by a cross power criteria a plurality of geometric source separation algorithms are obtained. To address the permutation, convolution and more general degrees of freedom for the shaping of beams, geometric information is utilized. The advantages of
10   bllind source separation and geometric beamforming are combined by minimizing the cross power spectra for multiple times t while enforcing constraints used in conventional adaptive beamforming.

BRIEF DESCRIPTION OF THE DRAWINGS

15   Figure 1 depicts a system for executing a software implementation of the present invention;

Figure 2 is a flow diagram of a method of the present invention;

DETAILED DESCRIPTION OF THE INVENTION

20

The present invention estimates values of $W(\omega)$ obtained by known blind source separation techniques by making geometric assumptions regarding the signal sources. The sources are assumed to be localized at least up to the spatial resolution of a given array. The invention assumes that signals originate from the same locations for the entire frequency
25   spectrum, permitting the formulation of geometric constraints on the filter coefficients. Different geometric constraints are introduced leading to a class of Geometric Source Separation algorithms.

Figure 1 depicts a system 100 for implementing the source separation method of the
30   present invention. The system 100 comprises a composite signal source 126 that supplies the signal that is to be separated into its component signals and a computer system 108 that executes the geometric source separation routine of the present invention. The source 126 may contain any source of composite signals, but is illustratively shown to contain a sensor

array 102, a signal processor 104 and a recorded signal source 106. The sensor array contains one or more transducers 102A, 102B and 102C such as microphones. The transducers are coupled to a signal processor 104 that performs signal digitization. A digital signal is coupled to the computer system 108 for signal separation and further processing. A recorded signal source 106 may optionally form a source of the compositie signals that require separation.

The computer system 108 comprises a central processing unit (CPU) 114, a memory 122, support circuits 116 and an input/output (I/O) interface 120. The computer system 108 is generally coupled through the I/O interface 120 to a display 112 and various input devices 110 such as a mouse and keyboard. The support circuits generally contain well known circuits such as cache, power supplies, clock circuits, a communications bus and the like. The memory 122 may include random access memory (RAM), read only memory (ROM), disk drive, tape drive and the like, or some combination of memory devices.

The invention is implemented as the geometric source separation routine 124 that is stored in memory 122 and executed by the CPU 114 to process the signal from the signal sources 126. As such the computer system 108 is a general purpose computer system that becomes a specific purpose computer system when executing the routine 124 of the present invention. Although a general purpose computer is illustratively shown as a platform for implementing the invention, the invention can also be implemented in hardware as an application specific integrated circuit (ASIC), a digital signal processing (DSP) integrated circuit, or other hardware device or devices. As such, the invention may be implemented in software, hardware or a combination of software and hardware.

Figure 2 depicts a flow diagram of the geometric source separation routine 124 of the present invention. At step 200, the composite, mixed signal 126 is input, the signal being separated into a plurality of data frames containing data samples of the input signal x(t). The routine 200 produces discrete Fourier transform (DFT) values x(ω) for each data frame x(t), that is, one frequency domain DFT value for each window of length T samples.

At step 204, a running time estimate of $R_{yy}(t,\omega)$ is computed from the outputs y(t). As large T are required for typical filter sizes, simultaneous diagonalization of $R_{yy}(t,\omega)$ for

multiple times t is performed so as to minimize the sum of squares of the off diagonal elements according to the equation 213:

$$J(W) = \sum_{t,w} \alpha(\omega) \parallel R_{yy}(t,\omega) - \text{diag}(R_{yy}(t,\omega)) \parallel^2 ,$$

with the factorization $R_{yy}(t,\omega) \approx W(\omega) R_{xx}(t,\omega) W^H(\omega)$. $\parallel ... \parallel$ refers to the Frobenius norm defined as $\parallel M \parallel^2 = \text{Tr}(MM^H)$. The summations over t and ω will range over all time and all frequency bins for which adaptation of W will occur, respectively. For faster convergence using gradient descent the total input power per frequency $\alpha(\omega)$ is normalized,

$$\alpha(\omega) = \sum_t \parallel R_{xx}(t,\omega) \parallel^{-2}.$$

This criteria is minimized with respect to the filter coefficients W. The lower bound of zero is obtained only if $R_{yy}(t,\omega)$ is diagonal.

The signal sources $s(t)$ are localized at angles $\theta = [\theta_1 ... \theta_m]$ and at a sufficient distance from the sensors 102 for a far field approximation to apply. While full three dimensional source location can be advantageously used with the present invention, the specific embodiment described here will identify souce locations simply by incident angle to the microphone array. Step 208 computes the geometric source locations based on the cross power spectra produced at step 204. Step 219 applies a selection or switch to determine which of various criteria will be next be applied to the cross power spectra calculation. For example, if the hard constraint 210

$$\text{diag}(W(\omega) D(\omega, \theta)) = 1$$

is applied to each filter $w_i(\omega)$, then the ith row vector in $W(\omega)$ is forced to have a unit response in the direction θ_i. In a gradient descent algorithm constraint 210 can simply be implemented by projecting the gradient $(\delta J/\delta w_i(\omega))$ to the linear subspace of permissible solutions with a constrained gradient. Since power or cross power minimization will try to minimize the response at the interference angles, this will lead to an equivalent singularity at

those frequencies. In those cases soft constraint 210 should be selected as a regularization term 211 of the form

$$J_1(\omega) = \left\| \ \text{diag} \ (W(\omega) \, D(\omega, \theta)) - 1 \ \right\|^{\ 2}.$$

5

A second, more resrictive hard the constraint 209

$$W(\omega) \, D(\omega, \theta) = I$$

10   may be selected a step 219. Hard constraint 209 imposes the conditions of hard constraint 210 to each filter $w_i(\omega)$ by forcing the ith row vector in $W(\omega)$ to have a unit response in the direction θ_i. Further, the hard constraint 209 requires that the ith row vector have a zero response in the direction of interfering signals θ_j where $i \neq j$.

15   At those frequencies where the grating lobes (periodic replica of the main lobes due to limited spatial sampling) of a beam pattern cross the interfering angles, $D(\omega, \theta)$ is not invertible. In those cases it is unreasonable to try to enforce constraint 209 as a hard constraint. Rather, soft constraint 212 is selected by adding the regularization term of the form

20

$$J_2(\omega) = \left\| \ (W(\omega) \, D(\omega, \theta)) - I \ \right\|^{\ 2}.$$

Step 219 performs various intialization conditions on equation 213. In all initializations the relationship $w_i(\omega)e_i = 1$ is required to normalize the scale during optimization. In step 215 the

25   filter structure is initialized to correspond to a delay-sum beamformer pointing into the orientations of the individual sources (del-sum). Using the orientations θ_i the filter coefficients of the ith beam (the row vectors $w_i(\omega)$ of $W(\omega)$) are initialized with algorithm

$$w_i(\omega) = \ \mathbf{d}(\omega, \theta_i)^H.$$

30

In step 219 the equation 213 can alternatively be initialized with beams that place zeroes at all orientations of interfering sources, that is, at angles θ/θ_i for the ith beam. The

initialization filters with minimum norm that satisify those conditions can be computed explicitly with a least squares approach resulting in the initialization algorithm 216,

$$w_i(\omega) = e_1^H \ [e_i, D(\omega, \theta/\theta_i)]^+ \ ,$$

where $^+$ indicates the psuedo-inverse and $[e_i, D(\omega, \theta/\theta_i)]$ is a matrix containing the ith column e_i of the unit matrix and all but the ith column of $D(\omega, \theta)$.

In an on line implementation of a separation algorithm the concept of introducing geometric information through an initialization is typically not feasible as the source positions in the environment may be changing dynamically. More frequently the geometric constraints are enforced at initialization and throughout the optimization process. The constraints change dynamically as the estimated locations change. The linear constraints 209 and 210 are each typically implemented as a soft constraint with a penalty term. The problem of noninvertibility is addressed by introducing a frequency dependent weighting of the penalty term. In particular the goal is to eliminate the constraints from the optimization for those frequency bands for which $D(\omega, \theta)$ is not inveritble. A rather straightforward metric for invertibility is the condition number. Therefore at steps 211 and 212 the regularization term $J(W)$ is weighted with the the inverse of the condition number of $\lambda(\omega) = \text{cond}^{-1}(D(\omega, \theta))$, which converges to zero when $D(\omega, \theta)$ is not invertible and remains bounded otherwise, for example as $0 \le \lambda(\omega) \le 1$. The total cost function 218 including frequency dependent weighting of the geometric regularization term is given by

$$J(W) + \lambda \sum_w \ \lambda(\omega) J_{2/1} \ W(\omega).$$

In algorithm 210 the regularization term J_1 will try to maintain the response of filters i in orientation θ_i. The delay-sum beamformer 215 satisfies the condition of algorithm 210 strictly. In algorithm 209 the regularization term J_2 will in addition minimize the response for the orientations of the interfering sources. The filter structure that strictly guarantees the constraints of algorithm 209 are computed with a least squares approach as the psuedo inverse of $D^H(\omega, \theta)$, or by including a regularization term βI for the noninvertibility problem, resulting in algorithm 219:

$$W(\omega) = D^H(\omega, \theta)\, (D(\omega, \theta)\, D^H(\omega, \theta) + \beta I)^{-1\cdot}$$

Algorithm 217 places a zero at the angles of interfering sources, while its response in other directions is not specified. The results for algorithms 215, 210 and 209 exhibit a main lobe in the directions of the corresponding source. For conflicting frequency bands, where a grating lobe coincides with the location of an interfering source, multiple cross power minimization cancels the main lobe for algorithm 215, while conserving it somewhat for algorithms 210 and 209 due to the geometric penalty. Qualitatively, the results for the data independent algorithm 219 appears to capture both main lobe and zeros in the correct location. However, its performance is inferior to the data adaptive algorithms 209, 210, 215 and 217.

Any of the algorithms 209, 210, 211 and 212 can be used to optimize the signal filter 222 and minimize cross power at step 202 using a gradient descent algorithm.. The angles θ_i of the multiple sources can be identified automatically using the Multiple Signal Classification (MUSIC) algorithm, which is a method of computing directions of arrival of multiple signals using arbitrarily placed antennas with arbitrary directional responses. The MUSIC algorithm also applies to measuring frequencies of multiple sinusoids comprising a sampled data time series. When used to measure frequency it is directly suited to tapped delay line implementations. The number of sources M is assumed to be known. While the examples and algorithms given in this description are directed toward numerical operations in the frequency domain, the present invention can also be implemented in the time domain.

CLAIMS

1. An apparatus for separating a convolutively mixed signal into a plurality of sources, comprising:

a plurality of signal sensors (102), each sensor residing at a sensor location;

a multiple input, multiple output signal filter (222);

means (204) for estimating cross powers of multiple output channels at multiple times;

means (208) for computing a desired set of source locations r, the source locations r being defined by spatial relation to the sensor locations; and

means for adapting the signal filter to minimize multiple output powers while enforcing a desired filter response for the desired set of source locations.

2. The apparatus of claim 1 wherein the cross powers are minimized by estimating and minimizing multiple cross power spectra of the signal filter outputs in a frequency domain.

3. The apparatus of claim 1 wherein the desired filter response is enforced as a linear constraint on the signal filter when the signal filter is minimizing multiple output powers.

4. The apparatus of claim 1 wherein the desired filter response is enforced by adding a regularization term when the signal filter is minimizing multiple output powers.

5. The apparatus of claim 2 wherein a unit response for source locations r are enforced as
$$\text{diag}\ (W(\omega)\,D(\omega, r\,)) = 1$$

6. The apparatus of claim 2 wherein a unit response and a zero response for source locations r are enforced as
$$W(\omega)\,D(\omega, r) = I$$

7. The apparatus of claim 1 wherein the signal filter (222) is a finite impulse response filter.

8. The apparatus of claim 2 wherein cross power minimization is implemented by minimizing
a cost function (213) defined as

$$J(W) = \sum_{t,w} \alpha(\omega) \; \| \; R_{yy}(t,\omega) - \mathrm{diag}(R_{yy}(t,\omega)) \; \|^{\;2} .$$

9. The apparatus of claim 2 wherein a unit response for source locations r is enforced with an
additional regularization term (211) defined as

$$J_1(W) = \; \| \; \mathrm{diag}\,(W(\omega)\,D(\omega,r)) - 1 \|^{\;2} .$$

10. The apparatus of claim 2 wherein a unit response and a zero response for locations r are
enforced with an additional regularization term (212) defined as

$$J_2(W) = \; \| \; (W(\omega)\,D(\omega,r)) - \mathbf{I} \; \|^{\;2} .$$

11. The apparatus of claim 1 wherein the source locations r are specified by angles relative
to a signal sensor location.

12. The apparatus of claim 1 wherein the signal filter (222) is initialized to behave as a beam
former filter.

13. The apparatus of claim 12 wherein the initialized signal filter represents a set of delay
sum beamformers aiming at the desired set of source locations r.

14. The apparatus of claim 1 wherein the source locations r are estimated from cross power
spectra.

15. A computer system (108) for processing nonstationary signals, comprising:

    at least one signal input port;

5   a central processing unit (114) configured to receive a composite signal from the
    signal input port;

    a memory (122) interconnected to the central processing unit (114); and

    a geometric source separation module (124) interconnected to the memory (122) to
process digitized signals stored in the memory.

10

16. The computer system of claim 15, wherein the geometric source separation module
    further comprises at least one finite impulse response filter (222), the composite
    signal being applied to the filter.

15 17. The computer system of claim 16, wherein the memory (122) further comprises
    spatial location data of sources of the composite signal, said spatial location data
    being coupled to the finite impulse response filter (222).

18. The computer system of claim 17, wherein the geometric source separation module
20   further comprises a plurality of algorithms (209, 210, 215, 217) adapted to
    compute a plurality of filter coefficients for the finite impulse response filter.

19. A method of separating a composite signal into a plurality of component signals
    comprising the steps of:

25   computing the cross power spectra $R_{yy}(t,\omega)$ of the plurality of component signals;
    simultaneously diagonalizing $R_{yy}(t,\omega)$ with an algorithm (213);
    applying a linear constraint (209,210) to the algorithm (213) to create filter
    coefficients; and
    filtering the composite signal based on the filter coefficients.

20. The method of claim 19, further comprising the step of defining the diagonalizing algorithm (213) as

$$J(W) = \sum_{t,w} \alpha(\omega) \| R_{yy}(t,\omega) - \text{diag}(R_{yy}(t,\omega)) \|^2 ,$$

with the factorization $R_{yy}(t,\omega) \approx W(\omega) R_{xx}(t,\omega) W^H(\omega).$

21. The method of claim 20, further comprising the step of choosing a linear constraint from a set of two linear constraints given by a first linear constraint (210)

$$\text{diag}(W(\omega) D(\omega, \theta)) = 1$$

and a second linear constraint (209)

$$(W(\omega) D(\omega, \theta)) = I .$$

22. The method of claim 21, further comprising the step of applying a regualarization term (211) to the diagonalizing algorithm (213) of

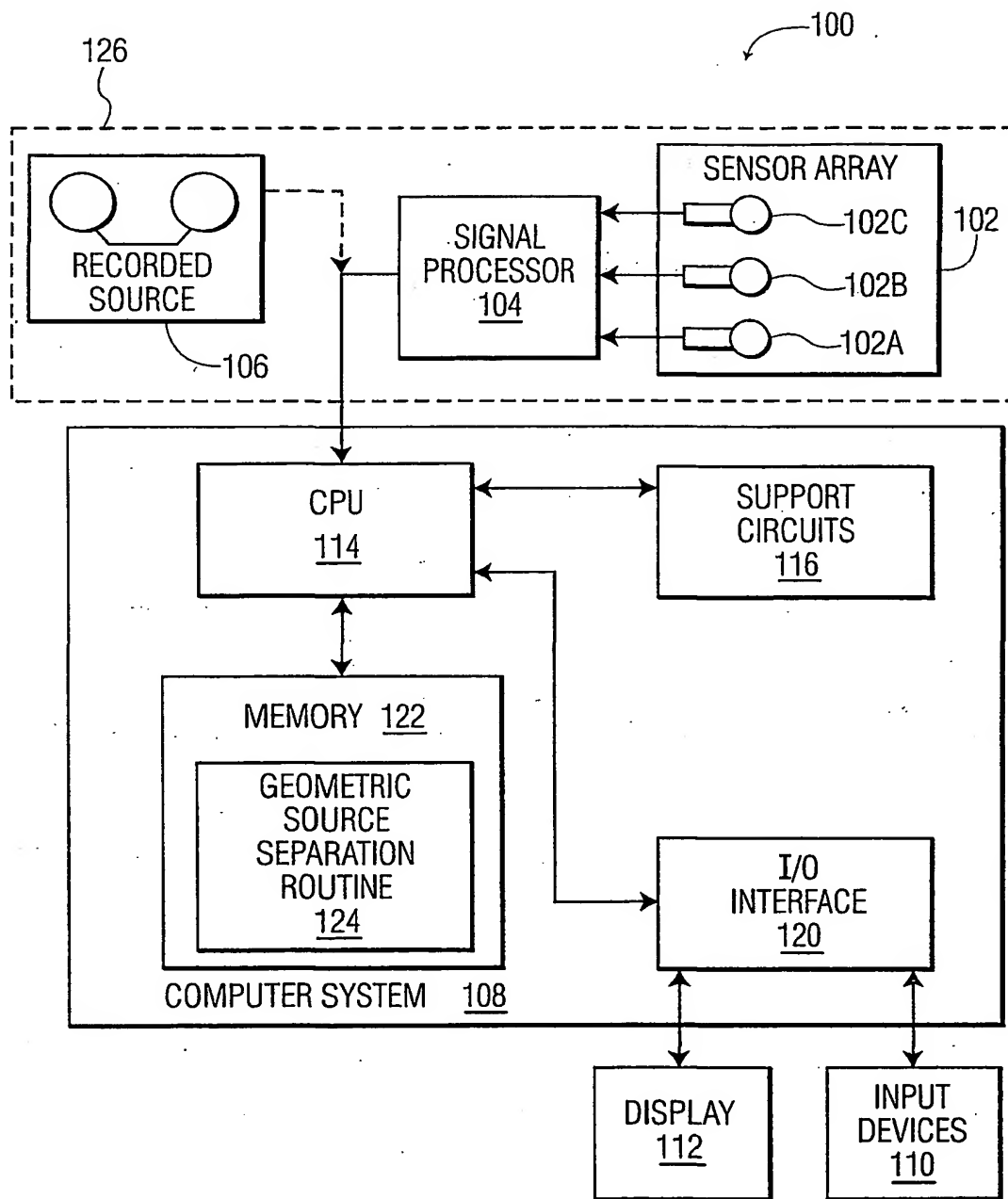$$J(\omega) = \| \text{diag}(W(\omega) D(\omega, \theta)) - 1 \|^2 .$$

FIG. 1

FIG. 2

INTERNATIONAL SEARCH REPORT
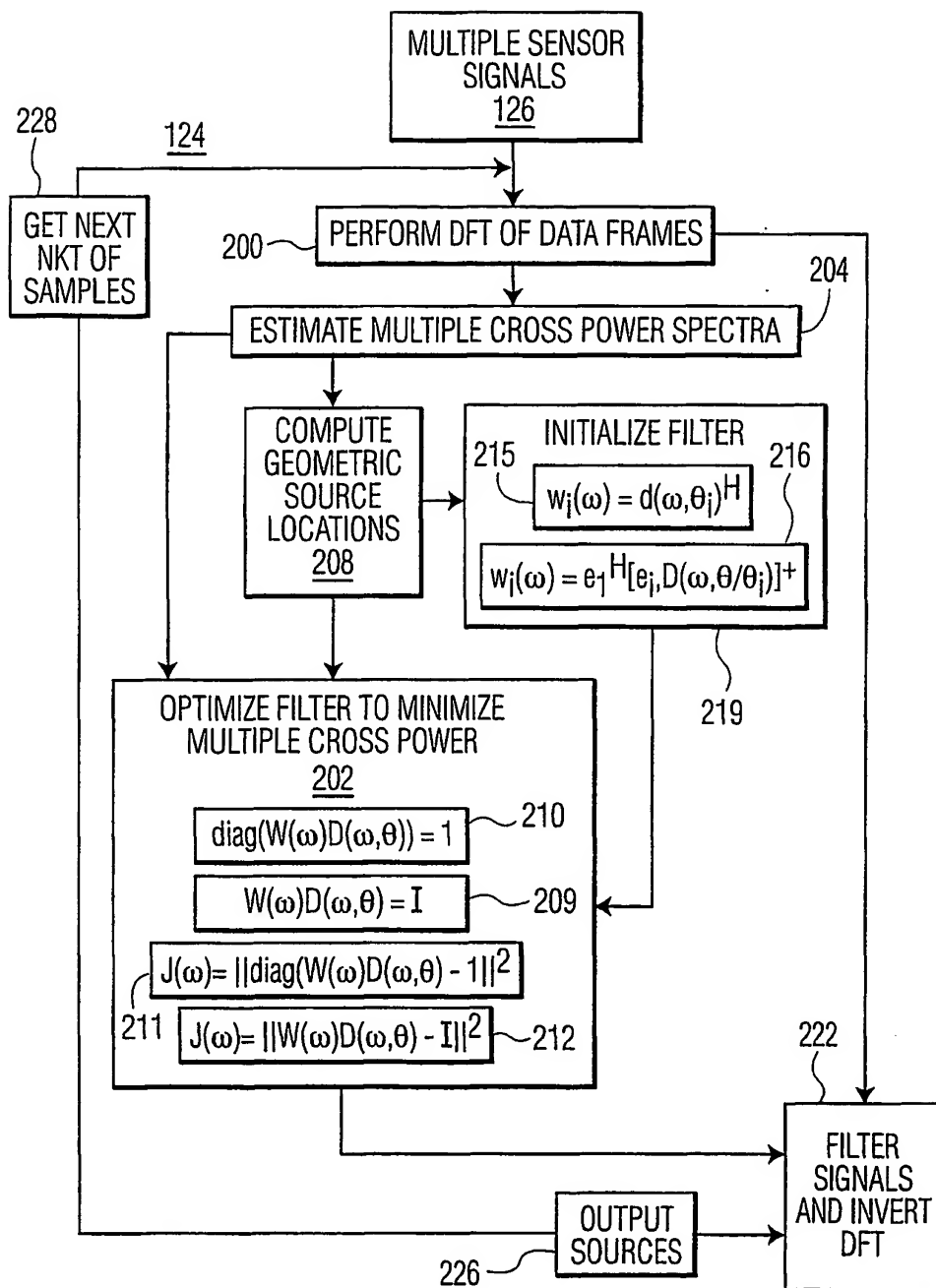
A. CLASSIFICATION OF SUBJECT MATTER

IPC 7   G10L21/02

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7   G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	AGEE B G: "Blind separation and capture of communication signals using a multitarget constant modulus beamformer" MILCOM 89. IEEE MILITARY COMMUNICATIONS CONFERENCE, BOSTON, MA, USA, vol. 2, 15 - 18 October 1989, pages 340-346, XP010083472 page 340, right-hand column, paragraph 2	1,3,12
X	WO 99 52211 A (SARNOFF CORP) 14 October 1999 (1999-10-14) cited in the application page 2, line 20 - line 27 page 3, line 6 - line 11 page 6, line 28 -page 9, line 9 page 10, line 8 - line 11	15,16,19
A		1-3,7,8, 11-14

—/—

[X] Further documents are listed in the continuation of box C.   [X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 June 2002	12/07/2002

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Ramos Sánchez, U

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	PARRA L ET AL: "Convolutive blind separation of non-stationary sources" IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, IEEE INC. NEW YORK, US, vol. 8, no. 3, May 2000 (2000-05), pages 320-327, XP002154443 ISSN: 1063-6676 sections II and III page 24 -page 28	20
P,X	PARRA L ET AL: "Geometric source separation: merging convolutive source separation with geometric beamforming" NEURAL NETWORKS FOR SIGNAL PROCESSING XI: PROCEEDINGS OF THE 2001 IEEE SIGNAL PROCESSING SOCIETY WORKSHOP (IEEE CAT. NO.01TH8584), NORTH FALMOU, pages 273-282, XP002203214 2001, Piscataway, NJ, USA, IEEE, USA ISBN: 0-7803-7196-8 the whole document	1-22

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 9952211	A	14-10-1999	US	6167417 A	26-12-2000
			AU	5265399 A	25-10-1999
			CN	1307747 T	08-08-2001
			EP	1070390 A1	24-01-2001
			JP	2002510930 T	09-04-2002
			WO	9952211 A1	14-10-1999